

Wright State University

CORE Scholar

---

International Symposium on Aviation  
Psychology - 2011

International Symposium on Aviation  
Psychology

---

2011

## The Effect of Measuring Situation Awareness on Pilot Behavior

J. van Eijck

C. Borst

M. Mulder

M.M. (René) van Paassen

Follow this and additional works at: [https://corescholar.libraries.wright.edu/isap\\_2011](https://corescholar.libraries.wright.edu/isap_2011)



Part of the [Other Psychiatry and Psychology Commons](#)

---

### Repository Citation

van Eijck, J., Borst, C., Mulder, M., & van Paassen, M. (2011). The Effect of Measuring Situation Awareness on Pilot Behavior. *16th International Symposium on Aviation Psychology*, 166-171.  
[https://corescholar.libraries.wright.edu/isap\\_2011/88](https://corescholar.libraries.wright.edu/isap_2011/88)

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2011 by an authorized administrator of CORE Scholar. For more information, please contact [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

# THE EFFECT OF MEASURING SITUATION AWARENESS ON PILOT BEHAVIOR

J. van Eijck, C. Borst, M. Mulder, M. M. (René) van Paassen  
Delft University of Technology, Faculty of Aerospace Engineering  
Kluijverweg 1, 2629HS Delft, The Netherlands

When developing new human-machine interfaces for aircraft cockpits, the contribution of the interface to the pilot's situation awareness is of interest because this might influence human decision-making. Besides subjective measures, a frequently used objective method of measuring SA is the freezing of the task, removing information from the displays and querying the pilot about the situation. Research has indicated that by doing so, the pilot's notion of what is important during this task may change, thereby potentially influencing decision-making. This paper describes an experiment in which pilots are given a terrain avoidance task, either with or without an interruption and SA query. The results showed a reorientation of the pilot's gaze without having a significant effect on the decision-making.

During the past decades, the number of automated systems in modern aircraft cockpits has increased significantly. Although the number of physical displays in the pilot's work environment decreased, the amount of information that needed to be presented on them increased. This has inevitably led to interfaces that are more complex and the question is often raised if the pilot is able to correctly interpret all the data that is presented. In other words: will the pilot still have sufficient Situation Awareness (SA)?

The concept of SA has been the subject of many discussions and therefore many different definitions have been developed (Uhlarik & Comerford, 2002). These definitions all seem to lack on some aspect of what is believed to be SA. Many definitions seem to neglect the dynamic nature of SA, while others provide no means of how to measure it. To further illustrate the disagreement about these different definitions: some researchers believe that more emphasis is required on the 'situation' aspect of SA (Flach, 1994). Others believe that the 'situation' part is well-defined, but 'awareness' requires more research (Sarter & Woods, 1991). A unified definition of SA appears to be far away, if not impossible. However, the most widely accepted definition of SA is the one by Endsley, who defines SA as *the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future* (Endsley, 1995). Although Endsley neglects to define what these meaningful elements are, she does provide the means to measure the SA using perception, comprehension and projection of the future status of the situation. Not only the definition of SA causes much discussion, but also its relation to human error is still not clear. Considering low SA as the cause of human error would only lead to circular reasoning (Flach, 1995). This also has implications on how to measure SA.

This paper describes an experiment that evaluates a variant of the freeze technique. The largest advantage of the freeze technique is that it is believed to be an objective method to assess the subject's SA. However, there is a risk of changing the subject's notion of what the 'meaningful elements' in a particular situation are. Therefore an experiment is performed to investigate the effect of the method on the subject's behavior. Similar research was performed in the past by McGowan and Banbury, but they tried to assess the effect of the method on the subject's SA, instead of behavior (McGowan & Banbury, 2004). Because there is no accepted measure for SA available, it is not possible to investigate the effect of the freeze technique on a subject's SA.

## Experiment

The focus of the experiment lies on investigating the effect of a query on the attention distribution of the pilot and the decisions that the pilot makes. The experiment is designed around a terrain avoidance task. To perform this task, pilots will use a Synthetic Vision Display (SVD) enhanced with overlays that depicted the aircraft maneuvering opportunities relative to the terrain (Borst, Mulder, & Paassen, 2010). The items of the SA query addressed the elements that could be perceived from the flight display.

## Subjects and instructions

A total of thirteen professional airline pilots participated in the experiment. With one of the pilots, the eye-tracker failed to produce reliable data, therefore data of only twelve pilots is used during the analysis. Half of the subjects was presented the SA query during the measurement runs, while the other subjects were not given this query.

The subjects were not briefed about the true goal of the experiment, as this might influence their behavior during the experiment. The pilots that were presented the SA query were given the task description to first observe the flight situation and avoid terrain obstacles. The pilots were instructed to select either a straight climb, climbing turn (left or right) or level turn (left or right) as escape maneuvers. Furthermore, the pilots were asked to deviate as little from their original course as possible. A straight climb was preferred above all other options and a level turn should only be made if no other solution is possible. Finally, pilots were not given any minimal separation with the terrain.

## Independent variables

The presence of the SA query (QUERY) was introduced as a between-subjects independent variable with two levels. In addition, since multiple scenarios were used for every pilot, the scenario (SCENE) was introduced as a within-subjects independent variable with ten levels.

## Dependent measures

**Decision-making** The pilots were instructed to select one of the five possible actions to avoid the terrain ahead of them and deviate as little as possible from their initial heading. They could either perform a straight climb, climbing turn (left or right) or a level turn (left or right). In addition, the pilot was instructed to deviate as little from the initial course as possible. This means that if a straight climb was possible to safely avoid the terrain ahead, this was considered the best option. With a level turn, the course deviation would be largest. This option should therefore only be chosen if no other possibility exists to avoid a collision with the terrain ahead. The quality of the decision was rated optimal or non-optimal. The decision taken by the pilot is of interest, because pilots that were presented a query may make different decisions than the ones that were not given this query. It is hypothesized that when the pilot is forced by the query to think through multiple options, this might have an effect on the resulting decision.

**Pilot gaze** A faceLAB eye-tracker is installed in the simulator for tracking the pilot gaze. The eye-tracker consists of a set of two cameras with infrared filters and three infrared sources. The infrared sources provide a constant illumination on the pilot's face, which can only be made visible with the infrared filters on the cameras. A schematic overview of the eye-tracker and screen set-up is given in Figure 1. Because it is known that attention is related to SA, the distribution of attention of the pilot is of interest as well. If the distribution of attention changes as a result of the SA query, this means that this method is influencing the SA that it intends to measure. The pilot gaze direction can be determined and combined with a model of the test environment. Although it is well known that it is dangerous to draw conclusions about direction of attention based on gaze direction, the SA query is hypothesized to force the pilot to consciously focus on certain aspects of the display to perform well on the query.

Eye-tracking needs to be done very accurately, because the part of the test environment that is of interest here is the screen with the interface, which contains a lot of information on a small surface area. The different aspects of the interface need to be distinguishable from each other. High accuracy with this eye-tracker is achievable when it is well calibrated. The different elements of interest are displayed in Figure 2. Most of the identified elements are static elements (e.g. altitude and speed indicators are always located on the same place), but some have a dynamic nature and these depend on the aircraft state at a specific point in time: the flight path vector and the expanding collision box. The calculated positions of these dynamic elements are continuously updated such that these data can be correlated with the output of the eye-tracker. The eye-tracker data considered in the analysis is the data obtained during the pre-freeze phase of every simulation run. If the query would affect the pilot's behavior, this would be best noticeable during the simulation phase before the query is presented. In addition, since the simulation runs do not have a fixed time span, this results in equally long measuring times for all runs.

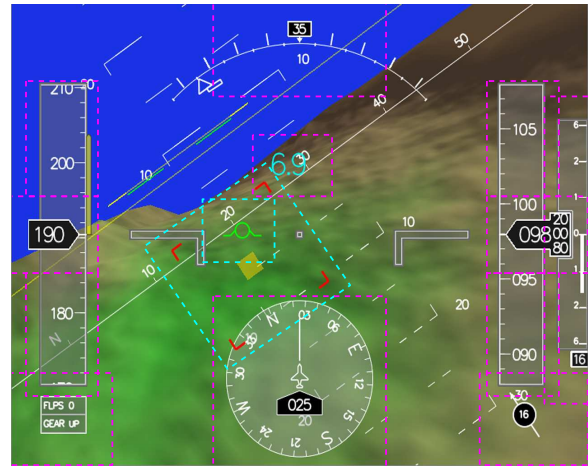
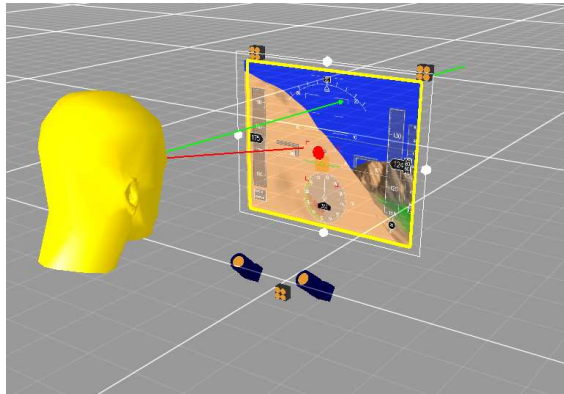


Figure 1: Test set-up with screen and eye-tracker.

Figure 2: Relevant elements for eye tracking.

In order to eliminate learning effects and individual differences from the data, the average time percentage per interface element over the last three training scenarios was used as a reference for gaze direction without the influence of a query. The same was done for the last three measurement runs and the difference between the two values is computed as a measure for change of gaze direction during the experiment. For the pilots without the query, this would be influenced by learning effects, while for the pilots that are given the query this change in gaze direction might have been caused by both learning effects and the effect of the SA query. By measuring the change of gaze direction instead of actual gaze direction in terms of time percentage per element, both individual differences and learning effects will be removed from the analysis.

The accuracy of the eye-tracker was evaluated between every four runs. This was done by means of a moving square on the screen that the pilot was requested to track. Both gaze-screen intersection coordinates and square position coordinates were logged and from the difference, the accuracy of the eye-tracker could be computed. When the eye-tracker accuracy during this test appeared to be of low quality, the subsequent simulation runs were discarded from the data.

**SA query performance** The effect of conditioning a pilot with the SA query on the performance on this query was also researched. During the last run of the experiment, also the pilots of the group without the query were given the SA query to see if they performed worse on answering these questions than the group of pilots that had been given this query during all the runs. It is expected that the latter will perform better on the query, although both groups have the same amount of experience with the display.

## Apparatus

The experiment was performed in the Human Machine Interaction (HMI) laboratory at the Faculty of Aerospace Engineering of the Delft University of Technology. The HMI-laboratory is a fixed-base flight simulator in a dark room. Two 18" LCD monitors were in front of the pilot, one showing the ESVD interface, the other showing the engine parameters. The aircraft model that was used during the simulation was a non-linear Cessna Citation 500 model. The aircraft was controlled using a control-loaded side stick on the right-hand side of the pilot and rudder pedals. The pilot was also able to adjust throttle and flap setting. The initial flap settings were randomly selected for every scenario as well as the wind conditions. Wind speed and direction remained constant during each run.

## Results

### Decision-making

The effect of QUERY on the pilot's decision-making was analyzed using Cochran's Q-test. There appeared to be no significant effect of QUERY on the decisions made by the pilot (Cochran,  $Q(1, 60) = 2.0, p = 0.16$ ). Figure 3 shows that pilots performing the experiment with a query took slightly less optimal decisions than pilots that were not given the query.

### Pilot gaze

The pilot gaze fixations on the elements defined in Figure 2 are expressed as a percentage of the thirteen seconds before performing a maneuver or filling in the query. The average gaze distribution during the last three training scenarios is used as the undisturbed gaze pattern for each pilot. The average gaze distribution during the last three measurement scenarios is influenced by learning effects and, for the group given the query, possibly the SA query. The questions in this query that test the SA level 1 of the pilots require them to remember values that are presented on the display. It is therefore hypothesized that the query will focus the pilot's gaze to the airspeed and altitude tapes, the compass rose and the wind speed and flaps/gear indicators.

The average gaze intersections for the sum of all these elements are calculated for both the training phase and measurement phase. Then, the difference between these two averages is calculated to see if the pilots who were given a query during the experiment have changed their gaze direction more than pilots that were not given this query. Based on these elements, the pilots with the query tended to adapt their gaze distribution significantly more during the experiment than the pilots without the query ( $U(1, 6) = 3.0, p = 0.016$ ). When decomposing this again in these different elements, it can be observed in Figure 4 that pilots from the query-group were on average focusing more on all of these elements, while Figure 5 shows that pilots who were not given this query tended to focus less on all of these elements.

The differences between training and measurement runs that are visible in Figures 4 and 5 were also analyzed using the Mann-Whitney test. There appeared to be no significant effect between training runs and measurement runs for any of the elements when no query is given to the pilot ( $U(1, 6) > 130.0, p > 0.30$ ), while the addition of a query results in significant effects for all of the elements ( $U(1, 6) < 90.0, p < 0.015$ ). These results indicate that the query significantly influences the pilot gaze direction. Screen shots with the gaze intersection points for two pilots of different groups in Figures 6 and 7 visualize the gaze reorientation due to the query.

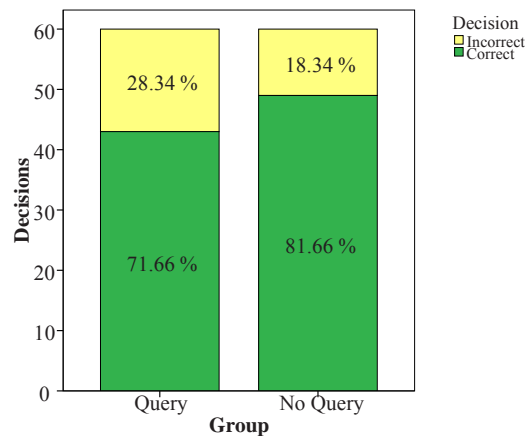


Figure 3: Decision-making per group of pilots.

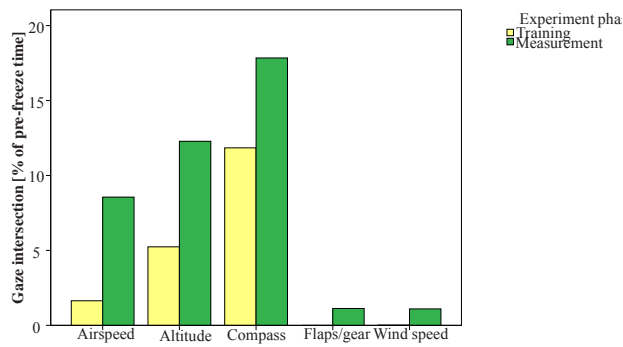


Figure 4: Gaze-screen intersection change during the experiment for pilots with the query.

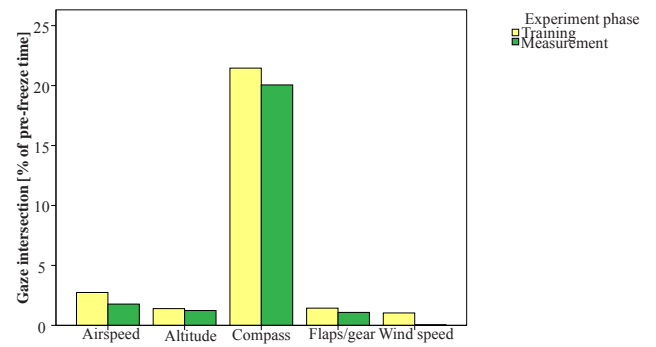
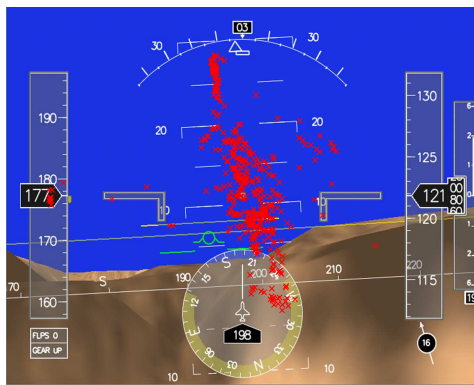
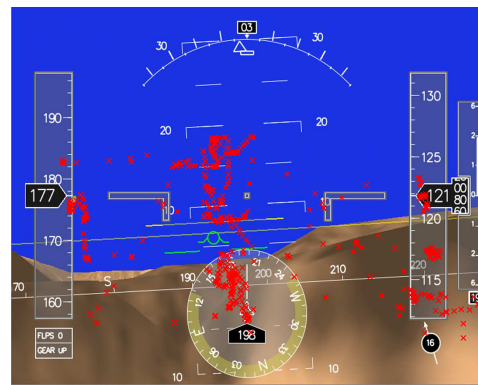


Figure 5: Gaze-screen intersection change during the experiment for pilots without the query.

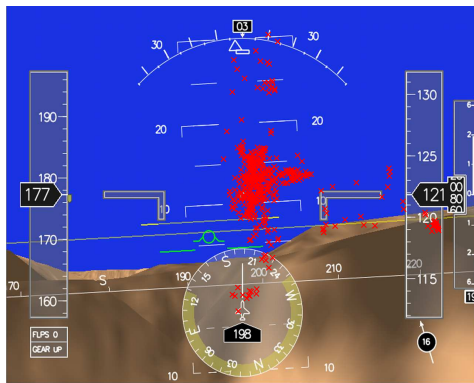


(a) During last training scenario

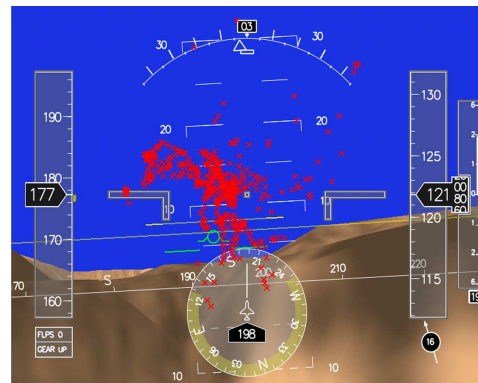


(b) During last measurement scenario

Figure 6: Gaze distributions for a pilot who was subjected to a query.



(a) During last training scenario



(b) During last measurement scenario

Figure 7: Gaze distributions for a pilot who was not subjected to a query.

### SA query performance

The scores on the SA query during the last measurement run showed no significant difference between the two groups of pilots ( $U(1, 6) = 18.0, p = 1.0$ ), but when decomposing the scores according to the levels of SA



by Endsley, some differences can be observed. The average partial and total SA scores are presented in Figure 8. Although not significantly ( $U(1, 6) = 6.5, p = 0.056$ ), the scores for SA level 1 are higher when the pilot knows what information is requested by the query. The query has a negative, but much smaller effect on scores for SA level 2 ( $U(1, 6) = 13.0, p = 0.42$ ) and SA level 3 ( $U(1, 6) = 9.5, p = 0.16$ ). The small positive effect of the query on SA level 1 is apparently compensated by small negative effects on SA level 2 and SA level 3, such that no effect can be observed on the total SA score.

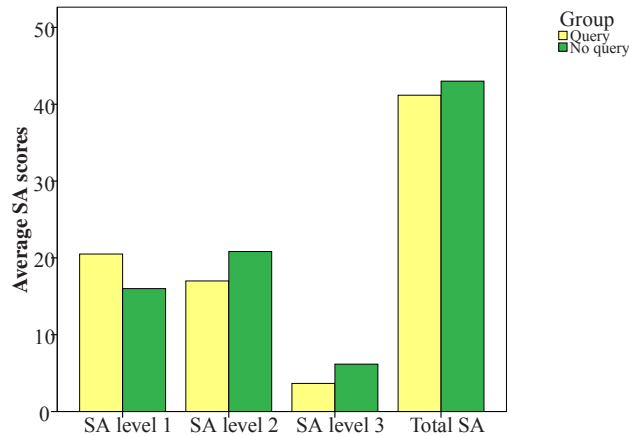


Figure 8: SA query scores for both groups of pilots.

## Conclusions

An experiment was conducted to investigate the possible intrusive effect of freeze techniques to measure situation awareness on the behavior of the pilot. The results indicated that pilots were indeed reorienting their gaze on the flight display to improve their performance on the SA query. Despite this reorientation, decision-making was not significantly affected as well as the total measured SA score. The conclusion we can draw here is that queries can be intrusive in terms of pilot scanning behavior, but also that they do not necessarily affect decision-making.

## References

- Borst, C., Mulder, M., & Paassen, M. M. van. (2010). Design and Simulator Evaluation of an Ecological Synthetic Vision Display. *Journal of Guidance, Control, and Dynamics*, 33(5), 1577–1591.
- Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, 37(1), 32-64.
- Flach, J. M. (1994). Situation Awareness: The Emperor's New Clothes. In M. Mouloua & R. Parasuraman (Eds.), *Human Performance in Automated Systems: Current Research and Trends*. Hillsdale.
- Flach, J. M. (1995). Situation awareness: Proceed with Caution. *Human Factors*, 37(1), 149-157.
- McGowan, A. M., & Banbury, S. P. (2004). Evaluating Interruption-Based Techniques using Embedded Measures of Driver Anticipation. In S. P. Banbury & T. S. (Eds.), *A Cognitive Approach to Situation Awareness* (p. 176-192). Ashgate.
- Sarter, N. B., & Woods, D. D. (1991). Situation Awareness: A Critical but Ill-Defined Phenomenon. *The International Journal of Aviation Psychology*, 1(1), 45-57.
- Uhlarik, J., & Comerford, D. A. (2002). *A Review of Situation Awareness Literature Relevant to Pilot Surveillance Functions* (Tech. Rep.). Manhattan KS Kansas State University.